



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Gene-environment interactions and colorectal cancer risk: an umbrella review of systematic reviews and meta-analyses of observational studies

Citation for published version:

Yang, T, Li, X, Montazeri, Z, Little, J, Farrington, S, Ioannidis, JPA, Dunlop, M, Campbell, H, Timofeeva, M & Theodoratou, E 2018, 'Gene-environment interactions and colorectal cancer risk: an umbrella review of systematic reviews and meta-analyses of observational studies', *International Journal of Cancer*.
<https://doi.org/10.1002/ijc.32057>

Digital Object Identifier (DOI):

[10.1002/ijc.32057](https://doi.org/10.1002/ijc.32057)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

International Journal of Cancer

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Gene–environment interactions and colorectal cancer risk: An umbrella review of systematic reviews and meta-analyses of observational studies

Tian Yang¹, Xue Li¹, Zahra Montazeri², Julian Little², Susan M. Farrington^{3,4}, John P.A. Ioannidis^{5,6}, Malcolm G. Dunlop^{3,4}, Harry Campbell¹, Maria Timofeeva¹ and Evropi Theodoratou^{1,4}

¹Centre for Global Health Research, Usher Institute of Population Health Sciences and Informatics, The University of Edinburgh, Edinburgh, United Kingdom

²School of Epidemiology and Public Health, University of Ottawa, Ottawa, Ontario, Canada

³Colon Cancer Genetics Group, Medical Research Council Human Genetics Unit, Medical Research Council Institute of Genetics & Molecular Medicine, Western General Hospital, The University of Edinburgh, Edinburgh, United Kingdom

Key words: colorectal cancer, diet, environment, gene, interaction, risk factor

Abbreviations: AICR: American Institute for Cancer Research; CI: confidence interval; CNKI: China National Knowledge Infrastructure; CRC: colorectal cancer; *CTNBI*: cadherin-associated protein β 1; CUP: continuous update project; *CYP24A1*: cytochrome P450 family 24 subfamily A member 1; GxE: gene–environment; *GATA3*: GATA binding protein 3; GECCO: Genetics and Epidemiology of Colorectal Cancer Consortium; GWAS: genome-wide association study; *HIATL1*: hippocampus abundant transcript-like 1; *MINK1*: misshapen like kinase 1; *NAT2*: N-acetyltransferase 2; NSAID: nonsteroidal anti-inflammatory drug; OR: odds ratio; *PI3K*: phosphatidylinositol-4,5-bisphosphonate 3-kinase; *PIK3C2G*: phosphatidylinositol-4-phosphate 3-kinase catalytic subunit type 2 gamma; *PTCHD3*: patched domain containing 3; RR: relative risk; *SHMT1*: serine hydroxymethyltransferase 1; SNP: single-nucleotide polymorphism; SOCCS: Scottish Colorectal Cancer Study; *TCF7L2*: transcription factor 7 like 2; WCRF: World Cancer Research Fund International

Additional Supporting Information may be found in the online version of this article.

M.T. and E.T. are joint and corresponding authors in this position

Grant sponsor: Cancer Research UK; **Grant numbers:** C31250/A22804C348/A18927; **Grant sponsor:** China Scholarship Council; **Grant sponsor:** Groupement des Entreprises Françaises dans la Lutte contre le Cancer; **Grant numbers:** PHRC-BRD09/C; **Grant sponsor:** National Cancer Institute; **Grant numbers:** CA137088P30, CA015704R01, CA143247R01201407U01, CA074778U01, CA122839U01/U24, CA074783U01/U24, CA074794U01/U24, CA074799U01/U24, CA074800U01/U24, CA074806U01/U24, CA097735U01, CA167551; **Grant sponsor:** National Heart, Lung, and Blood Institute; **Grant numbers:** HHSN268201100001C, HHSN268201100002C, HHSN268201100003C, HHSN268201100004C, HHSN268201100046C, HHSN271201100004C; **Grant sponsor:** National Institutes of Health; **Grant numbers:** K05 CA154337, K07 CA190673, U01 HG 004438, U01 HG004446, P01 CA033619, P01 CA055075, P01 CA087969, P50 CA127003, R01 CA042182, R01 CA063464, R01 CA076366, R01 CA137178, R01 CA151993, R01 CA48998, R01 CA60987, R35 CA197735, R35 CA197735, R37 CA54281, U01 CA074783, U01 CA167552, U01 CA167552, U01 CA186107, Z01 CP 010200; **Grant sponsor:** The Association Anne de Bretagne Génétique and the Ligue Régionale Contre le Cancer; **Grant numbers:** PHRC-BRD09/C; **Grant sponsor:** The California Department of Public Health; **Grant numbers:** HHSN261201000035C; **Grant sponsor:** The Canadian Institutes of Health Research; **Grant sponsor:** The German Research Council; **Grant numbers:** BR 1704/17-1, BR 1704/6-1, BR 1704/6-3, BR 1704/6-4, CH 117/1-1, HE 5998/2-1, HO 5117/2-1, KL 2354/3-1, RO 2270/8-1; **Grant sponsor:** The Groupement des Entreprises Françaises dans la Lutte contre le Cancer; **Grant numbers:** PHRC-BRD09/C; **Grant sponsor:** The Hawai'i Department of Health; **Grant numbers:** HHSN26120100037C, N01-PC- 67001, N01-PC-35137; **Grant sponsor:** The Interdisciplinary Research Program of the National Center for Tumor Diseases (NCT), Germany, and the German Federal Ministry of Education and Research; **Grant numbers:** 01ER0814, 01ER0815, 01ER1505A, 01ER1505B, 01KH0404; **Grant sponsor:** The Ontario Institute for Cancer Research; **Grant sponsor:** The Ontario Ministry of Research and Innovation; **Grant sponsor:** The Ontario Research Fund; **Grant sponsor:** The Regional Council of Pays de la Loire; **Grant numbers:** PHRC-BRD09/C; **Grant sponsor:** The Surveillance, Epidemiology and End Results (SEER) Program of the National Cancer Institute to Fred Hutchinson Cancer Research Center; **Grant numbers:** HHSN2612013000121, N01-CN-67009, N01-PC-35142; **Grant sponsor:** The Victoria Cancer Registry and Ontario Cancer Registry

DOI: 10.1002/ijc.32057

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

History: Received 15 Jun 2018; Accepted 15 Nov 2018; Online 11 Dec 2018

Correspondence to: Evropi Theodoratou, Usher Institute, The University of Edinburgh, Room 3.735, Doorway 1, Old Medical School, Teviot Place, Edinburgh EH8 9AG, United Kingdom, E-mail: e.theodoratou@ed.ac.uk; Tel.: +44(0)131 650 3210; or Maria N Timofeeva, Institute of Genetics and Molecular Medicine, The University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU, United Kingdom, E-mail: maria.timofeeva@igmm.ed.ac.uk; Tel.: +44 (0) 131 651 8631

⁴Cancer Research UK Edinburgh Centre, Medical Research Council Institute of Genetics & Molecular Medicine, Western General Hospital, The University of Edinburgh, Edinburgh, United Kingdom

⁵Stanford Prevention Research Center, Departments of Medicine, of Health Research and Policy, and of Biomedical Data Science, Stanford University School of Medicine, and Department of Statistics, Stanford University School of Humanities and Sciences, Stanford, California, USA

⁶Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, California, USA

The cause of colorectal cancer (CRC) is multifactorial, involving both genetic variants and environmental risk factors. We systematically searched the MEDLINE, EMBASE, China National Knowledge Infrastructure (CNKI) and Wanfang databases from inception to December 2016, to identify systematic reviews and meta-analyses of observational studies that investigated gene–environment (G×E) interactions in CRC risk. Then, we critically evaluated the cumulative evidence for the G×E interactions using an extension of the Human Genome Epidemiology Network's Venice criteria. Overall, 15 articles reporting systematic reviews of observational studies on 89 G×E interactions, 20 articles reporting meta-analyses of candidate gene- or single-nucleotide polymorphism-based studies on 521 G×E interactions, and 8 articles reporting 33 genome-wide G×E interaction analyses were identified. On the basis of prior and observed scores, only the interaction between rs6983267 (8q24) and aspirin use was found to have a moderate overall credibility score as well as main genetic and environmental effects. Though 5 other interactions were also found to have moderate evidence, these interaction effects were tenuous due to the lack of main genetic effects and/or environmental effects. We did not find highly convincing evidence for any interactions, but several associations were found to have moderate strength of evidence. Our conclusions are based on application of the Venice criteria which were designed to provide a conservative assessment of G×E interactions and thus do not include an evaluation of biological plausibility of an observed joint effect.

Introduction

Colorectal cancer (CRC) is the third most common cancer worldwide, with 746,000 new cases in men and 614,000 new cases in women.¹ In some low-to-middle-income countries, the incidence of CRC has been increasing partly due to changes in lifestyle and environment combined with aging populations.² Thus, it is critical to understand both modifiable and non-modifiable risk factors for CRC as this may enable more specific prevention strategies and risk assessment, especially in developing countries where CRC screening may not be feasible or affordable.³

CRC risk is determined by a complex interplay of both genetic variants and environmental exposures. It has been speculated that their interactions - known as gene–environment (G×E) interactions should also be important determinants of CRC risk. To date, genome-wide association study (GWAS) have shown that up to 50% of CRC heritability can be explained by common and rare variants included in popular genotyping arrays.⁴ Additional variants associated with CRC susceptibility that cannot be easily detected in GWAS by marginal effects of genetic factors may be identified by testing for interactions between single-nucleotide polymorphisms (SNPs) and environmental risk factors.^{5,6} Thus, identification of G×E interactions influencing CRC susceptibility may help to discover novel genetic and environmental risk factors for CRC, and extend our understanding of biological pathways and mechanisms of cancer etiology.

A number of systematic reviews, meta-analyses and genome-wide G×E interaction analyses that explored interaction effects in CRC have been published. We recently collected and evaluated the evidence across existing meta-analyses of observational studies in dietary factors and gene–diet interactions for

the 5 most common cancers.⁷ Here, we performed an umbrella review to collect, update, and assess the evidence across existing systematic reviews, meta-analyses and genome-wide G×E interaction analyses that have explored the joint effects between genes and a wider range of environmental exposures in CRC. Our aim is to provide an overview on the associations between G×E interactions and CRC risk and to pinpoint which of the associations have robust evidence by evaluating the strength of the evidence using predetermined guidelines.

Methods

Search strategy

We systematically searched the MEDLINE, EMBASE, China National Knowledge Infrastructure (CNKI) and Wanfang databases from inception to December 2016. The search strategy and Medical Subject Headings terms are displayed in Supporting Information Table S1. All identified publications went through a 2-step review before being included. Titles were reviewed by 1 investigator (TY). Abstracts and full texts were reviewed by 2 investigators (TY and MT). Any discrepancies were resolved by discussion.

Eligibility criteria

Three types of studies were eligible for this umbrella review: (i) systematic reviews of observational studies assessing interaction effects between genes and environmental exposures in CRC; (ii) meta-analyses of candidate gene- or SNP-based studies and analyses combining individual level data from multiple studies exploring G×E interactions and (iii) genome-wide investigation of G×E interactions on CRC risk within GWAS consortia. We excluded reviews without explicit systematic

literature searches; and systematic reviews or meta-analyses of observational studies that explored associations between CRC risk and genes or environmental exposures only.

Data extraction

One investigator (TY) extracted data which were then checked by a second investigator (ZM). For each eligible article, we extracted the first author, year of publication, the dietary and genetic risk factors examined, study design and the number of studies included. For meta-analyses, we extracted the summary study-specific relative risk estimates [relative risk (RR), odds ratio (OR)] along with the corresponding 95% confidence intervals (CIs), the number of cases and total participants, the *p* value for interaction and the *p* value (or *I*²) for heterogeneity.

Statistical analysis

For systematic reviews, we performed descriptive analyses and presented the authors' main conclusions. The evaluation process is described in detail in Supporting Information methods.

For meta-analyses of G×E interactions with a 2-sided *p* value for interaction <0.05, or for G×E interactions reaching genome-wide significance threshold, we used an extension of the Human Genome Epidemiology Network's Venice criteria^{8,9} to evaluate the strength of the evidence (Table 1^{10–17}). These guidelines have been used previously to assess cumulative evidence on joint effects of genes and environments on cancer risk.^{7,18}

First, we scored the strength of the observed evidence for the interaction between environmental exposures and genetic variants (observed score). Each G×E association was graded based on the *amount* of evidence, the extent of *replication* and the protection from *bias* (Table 1). On the basis of the combination of these 3 criteria, the epidemiological evidence for the association between G×E interaction and CRC risk was classified as strong, moderate or weak⁸ (Supporting Information Fig. S1).

Second, we established a prior score category (expected) for the G×E interactions using a framework presented in Boffetta *et al.*,⁸ which is based on prior scores for (i) the evidence of the main environmental and (ii) the evidence of the main genetic effects (Table 1). In brief, we established the prior score for the interactions based on the strength of evidence for the main environmental effect and the main genetic effect (1 = strong, 2 = moderate, 3 = weak) (Supporting Information Table S2). When both of the evidence were convincing (Class I), then the prior score category was strong. When one of the evidence was suggestive (Class III) or weak (Class IV), then the prior score category was weak.

Third, we examined the overall plausibility of each interaction by combining the prior score and the strength of the observed evidence. Higher weight was given to the observed evidence in case of conflicting results between the observed evidence and the prior scores.

Finally, for the statistically significant G×E interactions (with a 2-sided *p* value for interaction <0.05) that were identified from the candidate meta-analyses or for interactions that were concluded as suggestive by the authors of the systematic reviews, we tested the interactions in the Scottish Colorectal Cancer Study (SOCCS)¹⁹ dataset, and we also compared to results from the Genetics and Epidemiology of Colorectal Cancer Consortium (GECCO).²⁰

Results

Number and type of articles identified

Overall, 14,219 publications were identified across the 4 databases. After applying the inclusion and exclusion criteria, 42 publications were selected for inclusion (1 was in Chinese; Fig. 1). The details of 89 G×E interactions covering 22 environmental exposures identified in 15 systematic reviews of observational studies^{21–35} are presented in Supporting Information Results section.

Main findings of meta-analyses of candidate gene- or SNP-based studies

Twenty articles^{21,36–54} reporting meta-analyses of candidate gene- or SNP-based studies and analyses combining individual level data from multiple studies explored G×E interactions on CRC risk, covering 20 environmental exposures and 43 genes or genetic variants (Supporting Information Table S5). We identified 5 G×E interactions with 2-sided *p* value for interaction <0.05 (or adjusted *p* <0.05 after accounting for multiple comparisons): *N*-acetyltransferase 2 (*NAT2*) and processed meat intake⁵¹; *NAT2* and red meat intake⁵¹; rs16892766 (8q23.3) and vegetable consumption³⁹; serine hydroxymethyltransferase 1 (*SHMT1*) C1420T polymorphism and folate intake⁴⁶; and rs6983267 (8q24) and aspirin use⁴⁴ (Supporting Information Table S5). Also, the interactions between the above 5 environmental exposures (processed meat, red meat, vegetables, folate, aspirin use) and approximately 2.7 million genetic variants for CRC risk were also explored in GWAS consortia^{55–57} (Supporting Information Tables S6 and S7). However, none of the interactions observed in the candidate-based studies were detected at the genome-wide significance level in the GWAS consortia. We also tested interactions between rs16892766 and vegetable consumption, *SHMT1* C1420T polymorphism and folate intake and rs6983267 and aspirin use in the SOCCS,¹⁹ and compared to the results from the GECCO.²⁰ None of the associations were nominally significant ($\alpha = 0.05$) in our data (data not presented) or in the GECCO (data not presented). Each identified G×E interaction is described in detail in Supporting Information results.

Main findings of genome-wide investigation of G×E interactions within GWAS consortia

Eight articles^{55–62} corresponding to 33 genome-wide G×E interaction analyses explored joint effects between a large number of common polymorphisms and 22 selected

Table 1. Description of the extension of the Human Genome Epidemiology Network's Venice criteria that were used to assess cumulative evidence on joint effects of genes and environments on cancer risk

Steps	Description
Step 1 Score for the strength of the observed evidence for the G×E interactions	<p>First, we scored the strength of the observed evidence for the interaction between environmental exposures and genetic variants. Each G×E association was graded based on (i) the <i>amount of evidence</i>, (ii) the extent of <i>replication</i> and (iii) the <i>protection from bias</i>.</p> <p>(i) For the <i>amount of evidence</i>, the grade A, B or C was assigned when the total number of individuals in the smallest comparison group (assuming 1:1 ratio of cases and controls) in the meta-analysis was greater than 1,000, 100–1,000, or less than 100, respectively.</p> <p>(ii) The <i>replication consistency</i> was assessed by the reported heterogeneity: grade A, $I^2 < 25\%$; grade B, $25\% \leq I^2 \leq 50\%$; grade C, $I^2 > 50\%$ or p value for heterogeneity < 0.10.</p> <p>(iii) For <i>protection from bias</i> 3 aspects of G×E association were taken into account as suggested by Boffetta P <i>et al.</i>⁸: protection from bias for the environmental exposure, for the genetic analysis and for the overall interaction. Grade A means that bias, if present, may change the magnitude but not the presence of an association; grade B means that there is no evidence of bias that would invalidate an association, but important information is missing; and grade C means that there is a strong possibility of bias that would render the finding of an association invalid.</p> <p>On the basis of the combination of these 3 criteria (amount of evidence, degree of replication and protection from bias, each of which can be scored A, B and C), the epidemiological evidence for the association between G×E interaction and CRC risk was classified as strong, moderate or weak⁸ (Supporting Information Fig. S1).</p>
Step 2 Prior score (expected) for G×E interactions	<p>Second, we established a prior score category (expected) for the G×E interactions using a framework presented in Boffetta P <i>et al.</i>⁸, which is based on prior scores for (i) the evidence of the main environmental and (ii) the evidence of the main genetic effects (Supporting Information Table 2).</p> <p>(i) <i>Environmental main effect score</i>: We scored the main environmental effects based on the meta-analyses of the associations between environmental factors and CRC risk that were presented in the World Cancer Research Fund International (WCRF)/American Institute for Cancer Research (AICR) Third Expert Report,¹⁰ the subsequent Continuous Update Project (CUP) CRC reports¹¹ and the CUP CRC Systematic Literature Review 2016.¹² For the information of environmental risk factors that was not available in the above mentioned sections, we performed an additional literature search in MEDLINE and abstracted the relevant data as summarized and presented in Supporting Information methods. We then categorized the environmental factors in terms of strength of evidence by applying previously described set of criteria.¹³ The evidence was classified as convincing (Class I), highly suggestive (Class II), suggestive (Class III) or weak evidence (Class IV) based on sample size, highly significant summary associations, the 95% prediction intervals, presence of the small-study effect and the excess significance bias.</p> <p>(ii) <i>Genetic main effect score</i>: For the genetic main effects, a search in the National Human Genome Research Institute-European Bioinformatics Institute catalog of GWAS,¹⁴ the GWAS central database¹⁵ and MEDLINE was conducted as described in Supporting Information methods. Subsequently, we scored the genetic associations using the Human Genome Epidemiology Network Venice criteria.^{9,16,17} Only genetic effects with $p < 10^{-5}$ were considered for evaluation, and the evidence was classified as strong, moderate or weak based on a combination of the 3 criteria (amount of evidence, degree of replication and protection from bias), each of which was scored A, B or C (Supporting Information Fig. S1). For the genetic variants that reached genome-wide significance threshold, the evidence class of the genetic variant was only based on the amount of evidence.¹⁶ The search strategies, the Medical Subject Headings terms and the numbers of hits are presented in Supporting Information Tables 3 and 4, respectively.</p>
Step 3 Combined score	<p>Lastly, we examined the overall plausibility of each interaction by combining the prior score and the strength of the observed evidence. Higher weight was given to the observed evidence in case of conflicting results between the observed evidence and the prior scores.</p>

environmental exposures on CRC risk by using agnostic searches (a summary of all the analyses and details are presented in Supporting Information Tables S6; SNPs with the smallest p value for interaction from each genome-wide G×E interaction analysis are presented in Supporting Information Table S7). The following G×E interactions were identified that reached genome-wide significance threshold specified by the authors of the original publication: rs4143094 (10p14) and processed meat intake⁵⁵; rs9409565 (9q22.32) and light-to-

moderate alcohol drinking (1–28 g/day)⁶⁰; rs2965667 (12p12.3), rs16973225 (15q25.2) and aspirin and/or nonsteroidal anti-inflammatory drug (NSAID) use⁵⁷; patched domain containing 3 (*PTCHD3*) at 10p12.1, misshapen like kinase 1 (*MINK1*) at 17p13.2 and NSAID use⁶¹; rs964293 (20q13.2) and use of estrogen plus progestogen therapy⁵⁹; and rs1944511 (11q23.3) and overweight⁶² (Supporting Information Table S7). Each identified genome-wide G×E interaction is described in detail in Supporting Information results.

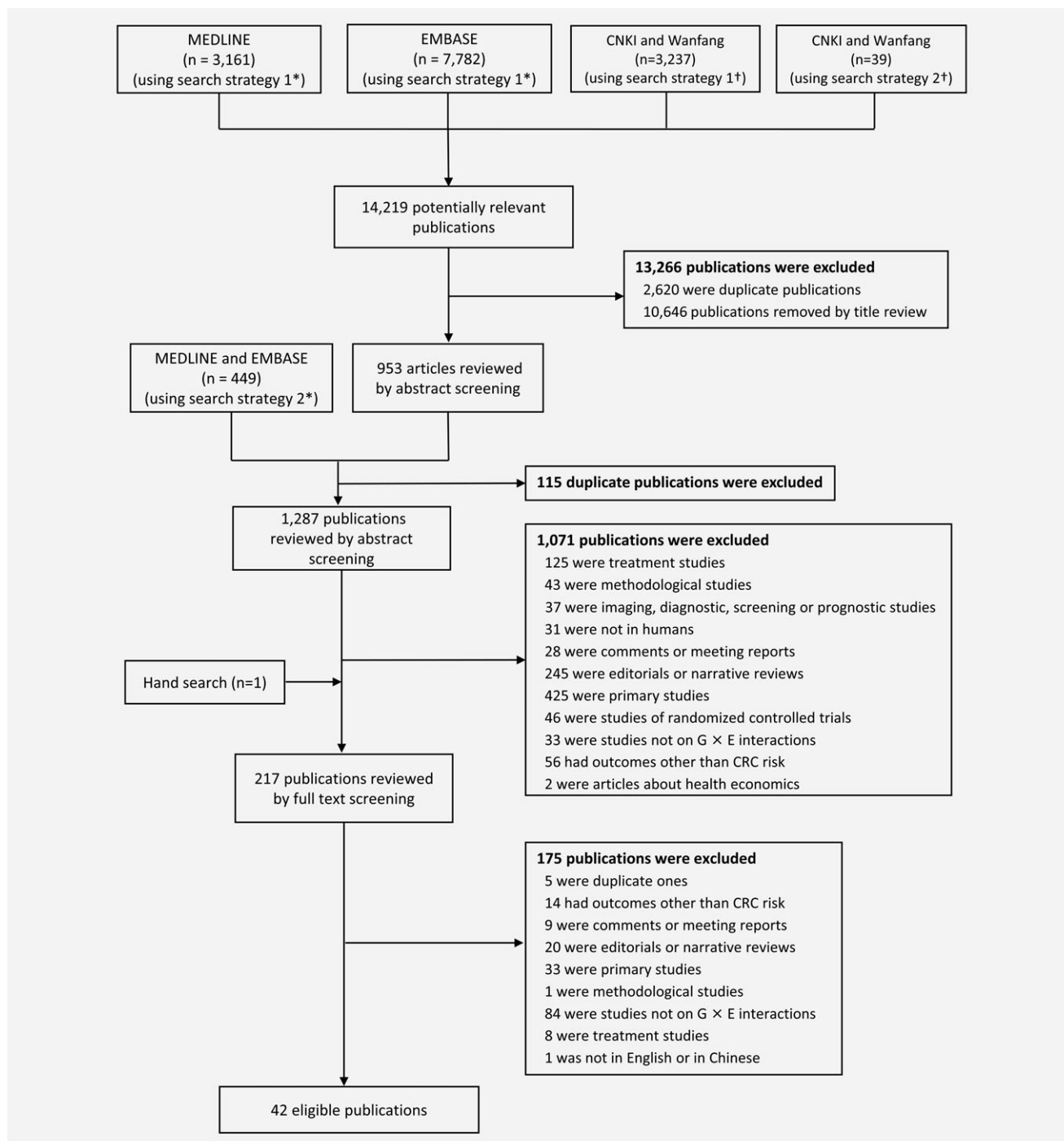


Figure 1. Flow chart of the literature search in MEDLINE, EMBASE, CNKI and Wanfang. *For the search in MEDLINE and EMBASE, we used both AND and OR to combine the keywords “G×E interactions” and “((gene* OR genom*) AND specific environmental risk factors)”, considering that there might be some publications that did not include the keyword “G×E interactions”. †For the search in CNKI and Wanfang, both strategies that included and not included specific environmental risk factors were used due to the limit of length of search strategies in these two databases.

Evaluation of the evidence for G×E interactions with main genetic effects of a $p < 10^{-5}$

Here, we present the evidence for the identified G×E interactions in relation to CRC risk with main genetic effects ($p < 10^{-5}$; Fig. 2).

Gene-aspirin use interactions. Aspirin use was associated with a reduced risk of CRC on the basis of a meta-analysis of 39 studies with 151,367 cases [users *versus* non-users RR, 0.79 (95% CI: 0.74, 0.85); $p = 7.8 \times 10^{-11}$; $I^2 = 91.1\%$],⁶³ thus the association was graded as highly suggestive (class II) due to

the high heterogeneity between the studies (Table 2). The main effect of rs6983267 (8q24) on CRC risk was graded as strong (ABA, equivalent to AAA based on the Venice criteria¹⁶) in a meta-analysis including 13,348 cases and 26,438 controls of European ancestry [OR, 0.84 (95% CI: 0.80, 0.88); $p = 7.45 \times 10^{-13}$; $I^2 = 37.7\%$]⁶⁴ (Table 3). Consequently, the interaction between rs6983267 (8q24) and aspirin use was given a moderate prior score (Moderate – 2) and a moderate overall plausibility score (Table 4).

Gene-vegetable interactions. A meta-analysis of the association between vegetable intake and CRC risk in the latest Continuous Update Project (CUP) CRC reports¹¹ and the CUP Colorectal Systematic Literature Review 2016¹² showed a reduction in CRC risk with 100 g/day increase in vegetable intake [RR, 0.98 (95% CI: 0.96, 0.99); $p = 0.01$; $I^2 = 0\%$; $n = 11$ prospective studies; n of cases = 14,136] and the association was classified as class IV (weak) (Table 2). On the basis of the strong (AAA, based on the Venice criteria¹⁶) main genetic (Table 3) and weak (class IV) environmental effects of vegetable intake on CRC risk, the possible interactions between 8q23.3 locus and vegetable intake on CRC risk was given a weak (Weak – 3) prior score and therefore no evidence was found for this interaction⁷ (Table 4).

Evaluation of the evidence for G×E interactions with no main genetic effects ($p > 10^{-5}$)

Here, we present the evidence for the identified G×E interactions in relation to CRC risk with no main genetic effects ($p > 10^{-5}$; Fig. 2). The G×E interactions were considered as tenuous even if they were classified as having moderate evidence.

Interactions between genetic variants and use of aspirin, NSAIDs or both. In a meta-analysis of 10 studies including 8,634 cases and 8,553 controls, regular use of aspirin and/or NSAIDs, compared to non-regular use, was associated with lower risk of CRC [RR 0.69 (95% CI: 0.64, 0.74); $p = 6.20 \times 10^{-28}$; $p_{\text{heterogeneity}} = 0.02$], thus the evidence was classified as highly suggestive (class II) due to the high heterogeneity between the studies⁵⁷ (Table 2). For the main genetic effects, no associations were observed between rs2965667 (12p12.3), rs16973225 (15q25.2) and CRC risk ($p > 10^{-5}$) in a meta-analysis of 7 GWAS from Europe including 8,749 cases and 18,245 controls ($p = 0.552$ and 0.242 , respectively)⁶⁵ (Table 3). Thus, the interactions between rs2965667 (12p12.3), rs16973225 (15q25.2) and aspirin and/or NSAID use were given moderate overall plausibility scores and weak (Weak – 3) prior scores (Table 4). Nevertheless, the overall plausibility scores for the interactions between 10p12.1/PTCHD3, 17p13.2/MINK1 and NSAID use could not be properly evaluated due to the missing elements of the extension of the Venice criteria⁸ that was used for assessing the observed evidence for the interactions (Table 4).

Interactions between genetic variants and use of estrogen plus progestogen therapy. The RR for use of estrogen plus

progestogen therapy on CRC risk was 0.74 (95% CI: 0.68, 0.81; $p < 0.001$; $I^2 = 0\%$) in a meta-analysis of 17 studies,⁶⁶ thus the association was classified as suggestive (class III; Table 2). Furthermore, no association was observed between the rs964293 variant and CRC risk ($p > 10^{-5}$) in the meta-analysis of 7 GWAS [OR, 0.97 (95% CI: 0.93, 1.01); $p = 0.156$; $I^2 = 6.3\%$]⁶⁵ (Table 3). On the basis of the prior and observed scores, the interaction between rs964293 (20q13.2) and use of estrogen plus progestogen therapy was given a moderate overall plausibility score and a weak (Weak – 3) prior score (Table 4).

Gene-alcohol interactions. A meta-analysis of 8 prospective studies in the latest CUP CRC reports¹¹ and the CUP Colorectal Systematic Literature Review 2016¹² showed one drink per day increase was associated with increased CRC risk [RR, 1.06 (95% CI: 1.00, 1.11); $p = 0.03$; $I^2 = 60.4\%$; n of cases = 36,942], and the association between light-to-moderate drinking and CRC risk was categorized as weak (class IV) (Table 2). Additionally, no main effect was observed for the rs9409565 (9q22.32) on CRC risk ($p > 10^{-5}$) of 18,299 cases [OR, 0.98 (95% CI: 0.95, 1.01); $p = 0.127$]⁶⁷ (Table 3). Hence, the interaction between rs9409565 (9q22.32) and light-to-moderate drinking was given a weak prior score (Weak – 3) and a moderate overall plausibility score (Table 4).

Gene-meat interactions. No evidence was found for the interactions between processed meat, red meat and NAT2 based on the weak (Weak – 3) prior score and the weak observed score (Tables 2–4). The possible 10p14 locus-processed meat interaction was given a weak prior score (Weak – 3) and a moderate plausibility score⁷ (Tables 2–4).

Gene-folate interactions. No evidence was found for the interaction between SHMT1 C1420T and folate intake on CRC risk based on a weak (Weak – 3) prior score and a weak observed score (Tables 2–4).

Gene-overweight interactions. No evidence was found for the interaction between rs1944511 (11q23.3) and overweight based on a weak (Weak – 3) prior score and a weak observed score (Tables 2–4).

Discussion

Main findings

Based on the prior and observed scores, only the interaction between rs6983267 (8q24) and aspirin use was found with a moderate overall plausibility score and a main genetic effect ($p = 7.45 \times 10^{-13}$; strong, based on the Venice criteria). In particular, the benefit of regular aspirin use on CRC risk was confined to individuals with T allele of rs6983267, which has been associated with impaired binding of cadherin-associated protein $\beta 1$ (CTNBN1)/transcription factor 7 like 2 (TCF7L2) and lower expression of MYC.^{68–70} Moreover, aspirin has been associated with Wnt pathway and the inhibition of nuclear

Table 2. Evaluation of environmental main effects for the environmental exposures identified in the selected GxE interactions

Environmental exposure	Reference	Unit of comparison	Number of cases	Number of studies	Relative risk (95% CI)	p Value	Prediction interval	Heterogeneity (I^2 and/or p value)	Evidence class ¹
Processed meat	WCRF, CUP 2017	Per 50 g/day	10,738	10	1.16 (1.08, 1.26)	0.0002	NA	$I^2 = 20.1\%$; $p = 0.258$	III
Red meat	WCRF, CUP 2017	Per 100 g/day	6,662	8	1.12 (1.00, 1.25)	0.05	NA	$I^2 = 23.6\%$; $p = 0.241$	IV
Light-to-moderate alcohol intake	WCRF, CUP 2017	1 drink/day	36,942	8	1.06 (1.00, 1.11)	0.03	NA	$I^2 = 60.4\%$; $p = 0.013$	IV
Vegetables	WCRF, CUP 2017	100 g/day	14,136	11	0.98 (0.96, 0.99)	0.01	NA	$I^2 = 0.0\%$; $p = 0.48$	IV
Total folate	WCRF, CUP 2017	100 mcg/day	4,633	8	0.99 (0.98, 1.00)	0.05	NA	$I^2 = 0.0\%$; $p = 0.92$	IV
Aspirin use	Qiao Y, 2018	Users vs. non-users	151,367	39	0.79 (0.74, 0.85)	7.8×10^{-11}	NA	$I^2 = 91.1\%$; $p = 0.000$	II ²
Aspirin and/or NSAID use	Nan H, 2015	Regular users of aspirin, NSAIDs or both vs. non-regular users	8,634	10	0.69 (0.64, 0.74)	6.20×10^{-28}	NA	$p = 0.02$	II ²
Use of estrogen plus progestogen therapy	Lin KJ, 2012	Ever users vs. non-users	NR	17	0.74 (0.68, 0.81)	<0.001	NA	$I^2 = 0\%$; $p = 0.88$	III
Overweight	WCRF, CUP 2017	Per 5 kg/m ² increase in BMI	71,089 (total number)	38 (20 for men and 24 for women)	Men: 1.08 (1.04, 1.11); women: 1.05 (1.02, 1.08)	Men: $p < 0.001$ Women: $p < 0.001$	NA	Men: $I^2 = 83.3\%$, $p < 0.001$; Women: $I^2 = 82.5\%$, $p < 0.001$	Men: III; Women: III

Abbreviations: BMI, body mass index; CI, confidence interval; CUP, Continuous Update Project; GxE, gene–environment; NA, not applicable; NR, not reported; NSAID, nonsteroidal anti-inflammatory drug; vs., versus; WCRF, World Cancer Research Fund.

¹Evidence class was decided using the after criteria: Convincing evidence (class I) required $>1,000$ cases, highly significant summary associations ($p < 10^{-6}$ by random effects), a 95% prediction interval not including the null, no evidence of small-study effects, no evidence of excess significance bias, and low heterogeneity values ($I^2 < 50\%$). Highly suggestive evidence (class II) required $>1,000$ cases, highly significant summary associations ($p < 10^{-6}$ by random effects), and the largest study to have a 95% confidence interval that excluded 1. Suggestive evidence (class III) required only $>1,000$ cases and $p < 0.001$ by random effects. Evidence was considered weak (class IV) when $p < 0.05$. No association indicates evidence for the main environmental effects with $p > 0.05$.

²The evidence was classified as highly suggestive (class I) due to the high heterogeneity between the studies.

Table 3. Evaluation of genetic evidence for variants identified in the selected G×E interactions

Genetic variant	Gene (or near gene)	Reference	Discovery sample size	Replication sample size / Number of studies in meta-analysis	Reported OR (95% CI)	p Value for genetic main effect	Heterogeneity, I^2	Venice criteria	Evidence class ¹
rs4143094	10p14/GATA3	Figueroa et al., 2014	9,287 cases and 9,117 controls of European ancestry from USA, Australia, Canada and Germany	Meta-analysis, 10 studies	NR	0.26	NR	NA	No association
Slow/intermediate/rapid	NAT2	Zhang L, 2012	13,606 cases and 17,957 controls of Africans, Asians, Caucasians and mixed populations	Meta-analysis, 39 studies	Slow vs. rapid phenotype: 0.96 (0.90, 1.01)	No association	$I^2 = 17.8\%$	NA	No association
Slow/intermediate/rapid	NAT2	Wang H, 2015	2,186 cases and 3,736 controls of Japanese; 466 cases and 4,356 controls of African Americans	Meta-analysis, 7 studies	Rapid vs. slow phenotype: Japanese: 1.05 (0.87, 1.27); African Americans: 0.75 (0.50, 1.14); Combined: 0.99 (0.83, 1.18)	Japanese: 0.77; African Americans: 0.19; Combined: 0.81	NR	NA	No association
rs9409565	9q22.32/HIATL1	Schumacher FR, 2015 ²	18,299 cases and 19,656 controls of European ancestry from North America, Australia and Europe	Meta-analysis, 4,725 cases and 9,969 controls of East Asian ancestry from Republic of Korea, China and Japan	0.98 (0.95, 1.01)	0.127	NR	NA	No association
rs16892766	8q23.3/EIF3H	Li M, 2015	41,728 cases and 44,393 controls	Meta-analysis, 11 studies	1.22 (1.18, 1.27)	1.39×10^{-24}	$I^2 = 4\%$	AAA	Strong
rs6983267	8q24.21	Tanskanen T, 2017	1,701 Finnish cases and 14,082 population-based, cancer-free controls	Meta-analysis, 13,348 cases and 26,438 controls of European ancestry	0.84 (0.80, 0.88)	7.45×10^{-13}	$I^2 = 37.7\%$	ABA (equivalent to AAA)	Strong
C1420T	SHMT1	Wang Q, 2014	3,912 cases and 4,954 controls	Meta-analysis, 7 studies	TT vs. CC: 0.84 (0.73, 0.97); CT vs. CC: 1.01 (0.92, 1.10); TT + CT vs. CC: 0.97 (0.89, 1.06); TT vs. CT + CC: 0.84 (0.73, 0.96)	TT vs. CC: 0.020; CT vs. CC: 0.903; TT + CT vs. CC: 0.476; TT vs. CT + CC: 0.013	TT vs. CC: $I^2 = 3.8\%$; CT vs. CC: $I^2 = 0\%$; TT + CT vs. CC: $I^2 = 0\%$; TT vs. CT + CC: $I^2 = 0\%$	NA	No association

(Continues)

Table 3. Evaluation of genetic evidence for variants identified in the selected G×E interactions (Continued)

Genetic variant	Gene (or near gene)	Reference	Discovery sample size	Replication sample size / Number of studies in meta-analysis	Reported OR (95% CI)	p Value for genetic main effect	Heterogeneity, I^2	Venice criteria	Evidence class ¹
rs2965667	12p12.3/ <i>PIK3C2G</i>	Orlando G, 2016	8,749 cases and 18,245 controls from Europe	Meta-analysis, 7 studies	0.97 (0.87, 1.08)	0.552	$I^2 = 4.8\%$	NA	No association
rs16973225	15q25.2/ <i>Interleukin 16</i>	Orlando G, 2016	8,749 cases and 18,245 controls from Europe	Meta-analysis, 7 studies	1.05 (0.97, 1.15)	0.242	$I^2 = 0\%$	NA	No association
rs964293	20q13.2/ <i>CYP24A1</i>	Orlando G, 2016	8,749 cases and 18,245 controls from Europe	Meta-analysis, 7 studies	0.97 (0.93, 1.01)	0.156	$I^2 = 6.3\%$	NA	No association
Including 7 variants ³	10p12.1/ <i>PTCHD3</i>	Timofeeva M, 2015	8,100 cases and 21,820 controls from Europe	Meta-analysis, 6 studies	NR	0.352	NR	NA	No association
Including 8 variants ⁴	17p13.2/ <i>MINK1</i>	Timofeeva M, 2015	8,100 cases and 21,820 controls from Europe	Meta-analysis, 6 studies	NR	0.381	NR	NA	No association
rs1944511	11q23.3	Siebert S, 2013	259 cases and 1,002 controls	Genome-wide G×E interaction analysis	1.07	0.536	NR	NA	No association

Abbreviations: CI, confidence interval; G×E, gene–environment; NA, not applicable; NR, not reported; OR, odds ratio; vs., versus.

¹Evidence class was decided on the basis of the Human Genome Epidemiology Network's Venice criteria: No association indicates evidence for main genetic effects with $p > 10^{-5}$. Only genetic effects with $p < 10^{-5}$ were considered for evaluation. On the basis of a combination of 3 criteria (amount of evidence, degree of replication, and protection from bias) each of which can be scored A, B, or C), the epidemiological evidence for an effect of the genotype is classified as strong, moderate, or weak. For amount of evidence, a grade of A, B, or C was assigned when the sample size for the rarer genotype in the meta-analysis was greater than 1,000, 100–1,000, or less than 100, respectively. For replication consistency, we used $P < 25\%$ to assign grade A, 25–50% to assign grade B, and $> 50\%$ or a p value for heterogeneity < 0.10 to assign grade C. For protection from bias, a grade of A means that bias, if present, may change the magnitude but not the presence of an association; a grade of B means that there is no evidence of bias that would invalidate an association, but important information is missing; and a grade of C means that there is a strong possibility of bias that would render the finding of an association invalid. For the genetic variants that reached genome-wide significance threshold, the evidence class of the genetic variant was only based on the amount of evidence based on the clarification of Venice Criteria (Khoury MJ *et al.*, 2009).

²Current study exploring the marginal association of rs9409565 was used since it is about twice as large as Gong *et al.* (2016).

³Seven variants at 10p12.1 were included in the analysis that explored main genetic effects by Timofeeva M *et al.*, 2015. However, the interaction analysis (by Iiao S *et al.*, 2015) included 8 variants: chr10:27687284, chr10:27687437, chr10:27687638, chr10:27687989, chr10:27688101, chr10:27702174 and chr10:27702624.

⁴Eight variants at 17p13.2 were included in the analysis that explored main genetic effects Timofeeva M *et al.*, 2015. However, the interaction analysis (by Iiao S *et al.*, 2015) included 4 variants: chr17:4794313, chr17:4794407, chr17:4796839 and chr17:4797910.

Table 4. Evaluation of evidence for the selected G×E interactions in relation to CRC risk

Environmental exposure	Genetic variant	Gene (or near gene)	Score based on observed evidence ¹	Strength of observed evidence for interaction	Score for environmental evidence (evidence class)	p Value for main genetic effect	Score for genetic evidence / Venice criteria ²	Prior score ³	Combined score ⁴
<i>Meta-analyses of candidate gene- or SNP-based studies</i>									
Processed meat	Slow/intermediate/rapid	MAT2	(B or C) BC	Weak	III	No association	No association	Weak: 3	No evidence
Red meat	Slow/intermediate/rapid	MAT2	(B or C) AC	Weak	IV	No association	No association	Weak: 3	No evidence
Vegetables	rs16892766	8q23.3/EIF3H	CBB	Weak	IV	1.39×10^{-24}	Strong/AAA	Weak: 3	No evidence
Folate intake	C1420T	SHMT1	CAC	Weak	IV	TT vs. CC: 0.020; CT vs. CC: 0.903; TT + CT vs. CC: 0.476; TT vs. CT + CC: 0.013	No association	Weak: 3	No evidence
Aspirin use	rs6983267	8q24	BAB ⁵	Moderate	II	7.45×10^{-13}	Strong/ABA (equivalent to AAA)	Moderate: 2	Moderate
<i>Genome-wide G×E interaction analyses</i>									
Processed meat	rs4143094	10p14/GATA3	BBB	Moderate	III	0.26	No association	Weak: 3	Moderate
Light-to-moderate drinking	rs9409565	9q22.32/HIAT1	BBA	Moderate	IV	0.127	No association	Weak: 3	Moderate
Aspirin and/or NSAID use	rs2965667	12p12.3/PIK3C2G	B - A	Moderate	II	0.552	No association	Weak: 3	Moderate
NSAID use	rs16973225	15q25.2/interleukin 16	B - A	Moderate	II	0.242	No association	Weak: 3	Moderate
	Including 8 variants	10p12.1/PTCHD3	- - B	Not possible to evaluate ⁶	II	0.352	No association	Weak: 3	Not possible to evaluate
	Including 4 variants	17p13.2/MINK1	- - B	Not possible to evaluate ⁶	II	0.381	No association	Weak: 3	Not possible to evaluate
Use of estrogen plus progestogen therapy	rs964293	20q13.2/CYP24A1	BBA	Moderate	III	0.156	No association	Weak: 3	Moderate
Overweight	rs1944511	11q23.3	C - B	Weak	III	0.536	No association	Weak: 3	No evidence

Abbreviation: CRC, colorectal cancer; G×E, gene–environment; NSAID, nonsteroidal anti-inflammatory drug; SNP, single-nucleotide polymorphisms; vs., versus.

¹The strength of the observed evidence for interaction between the environmental exposures and the genetic variants was based on an extension of the Human Genome Epidemiology Network's Venice criteria used for assessing cumulative evidence for genetic associations. Each G×E association was graded based on the amount of evidence, the extent of replication and protection of bias. Dashes indicate that 1, 2, or 3 elements of the Venice criteria cannot be decided. A complete score should have 3 letters, corresponding to amount of evidence, degree of replication, and protection from bias components of the Venice criteria. If 1 element is missing, the score is represented by a single dash and 2 letters. If 2 elements are missing, the score is represented by 2 dashes and a letter.

²No association indicates evidence for main genetic effects with $p > 10^{-5}$. Only genetic effects with $p < 10^{-5}$ were considered for evaluation using the Human Genome Epidemiology Network's Venice criteria.

³The prior score was based on scores for environmental evidence and genetic evidence (Table 2 and Table 3).

⁴The overall plausibility of an interaction was examined by comparing the prior score and the score for the strength of the observed evidence. Higher weight was given to the observed evidence in case of conflicting results between the prior and observed scores.

⁵The replication consistency was graded as A because the interaction between aspirin use and rs6983267 was replicated in the GECCO.

⁶Jiao *et al* (2015) presented no information on the total number of individuals in the smallest comparison group and heterogeneity between the studies. Therefore evaluating the amount of evidence and the extent of replication according to the Venice criteria was not possible.

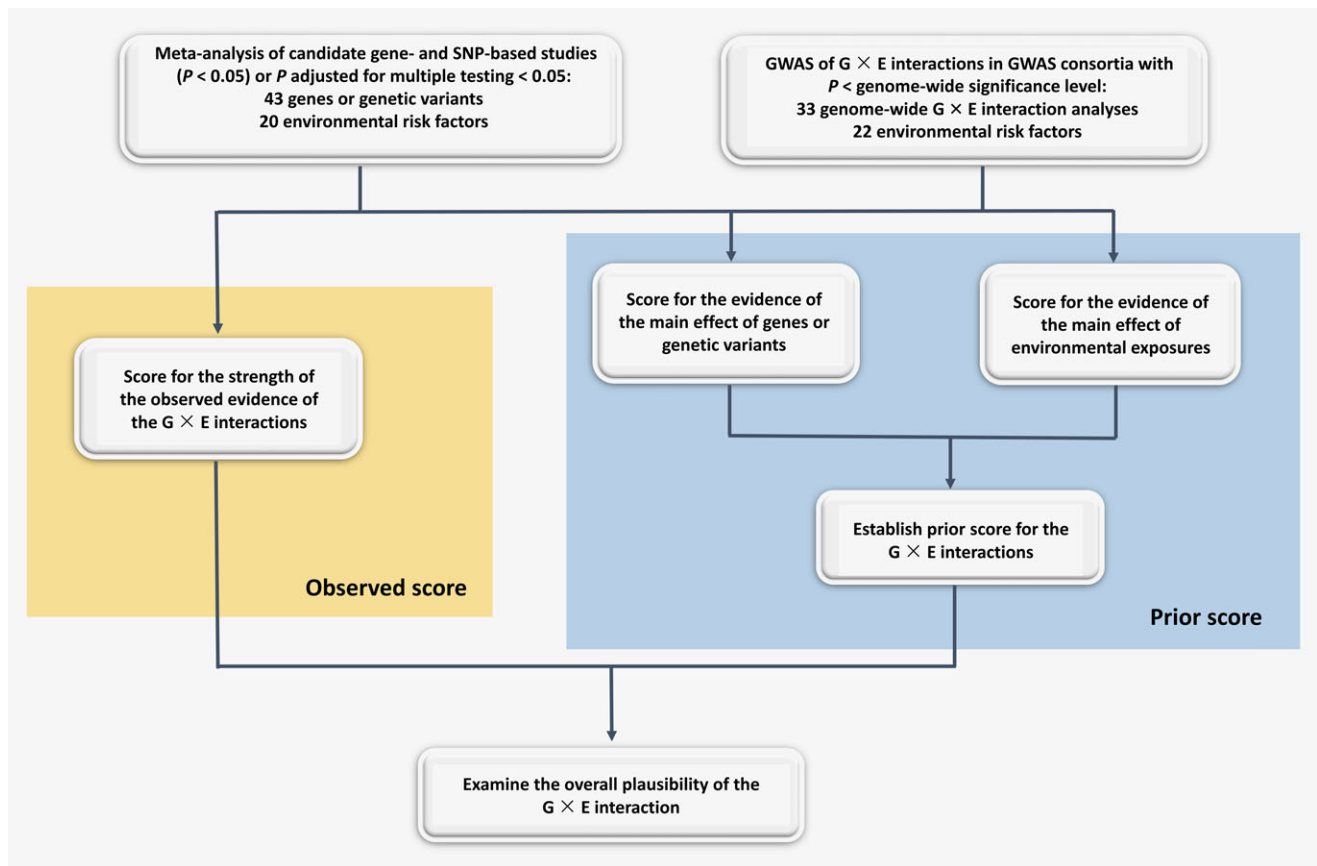


Figure 2. Steps in assessing GxEx interactions with P for interaction < 0.05 or reached genome-wide significance thresholds. [Color figure can be viewed at wileyonlinelibrary.com]

CTNNB1 expression in colon cancer cell lines.⁷¹ Hence, it is suggested that a genetic background by which *CTNNB1*/*TCF7L2* binding is not constitutively active is necessary for the susceptibility to the effects of aspirin on the Wnt/*CTNNB1* pathway.⁴⁴

Moderate overall plausibility scores were also found for the interactions between rs4143094 (10p14) and processed meat intake, rs9409565 (9q22.32) and light-to-moderate alcohol drinking (1–28 g/day), rs964293 (20q13.2) and use of estrogen plus progestogen therapy, as well as rs2965667 (12p12.3), rs16973225 (15q25.2) and aspirin and/or NSAID use. However, these interactions are regarded as tenuous due to the lack of main genetic effects ($p > 10^{-5}$), even though they may provide clues to discovering novel CRC susceptibility loci that have not been readily detected in GWAS by their marginal effects of genetic factors.^{5,6}

Little is known about the underlying molecular mechanisms of the interactions between rs2965667 at the 12p12.3/ phosphatidylinositol-4-phosphate 3-kinase catalytic subunit type 2 gamma (*PIK3C2G*) locus, rs16973225 (15q25.2/*interleukin 16*) and aspirin and/or NSAID use on CRC risk. *PIK3C2G* gene encodes a protein of the phosphatidylinositol-4,5-bisphosphonate 3-kinase (*PI3K*) family,⁷² of which the activated signaling can inhibit apoptosis in colon cancer cell lines that can be restored

with NSAID-mediated blockade of *PI3K*.^{57,73} *Interleukin 16* may stimulate monocyte induction of proinflammatory cytokines associated with tumorigenesis,^{74,75} which suggests that polymorphisms in or near the *interleukin 16* gene may be associated with the production of inflammatory cytokines that modify the chemopreventive effect of aspirin and/or NSAIDs on CRC.⁵⁷ However, it is proposed that those GWAS-identified promising loci outside of a known coding region may affect more distant genes rather than the closest gene.⁷⁶

The interaction between use of estrogen plus progestogen hormone preparations and the rs964293 variant in the 20q13.2, known as cytochrome P450 family 24 sub-family A member 1 (*CYP24A1*) is biologically plausible. *CYP24A1* is greatly expressed in malignant colon tumours,⁷⁷ and some variants in *CYP24A1* have been associated with CRC risk.⁷⁸ In a recent meta-analysis, ever use of estrogen plus progestogen therapy has been associated with lower CRC risk [OR: 0.74 (95% CI: 0.68, 0.81); $p < 0.001$].⁶⁶ Also, an inverse association was found in the meta-analysis of randomized controlled trials with a hazard ratio of 0.77 (95% CI: 0.59, 0.98) ($p = 0.037$).⁶⁶ It is suggested that *CYP24A1* may only be a metabolizing enzyme for progestogens but not estrogen, since an interaction effect was only found with the use of combined estrogen-progestogen therapy and not with estrogen-only intake.⁵⁹

The mechanism of the modifying effect of the rs4143094 variant at the 10p14 locus, near GATA binding protein 3 (*GATA3*) region, on the association between processed meat intake and CRC risk is even less clear. *GATA3* has been described as a master regulator of T-helper 2 cell differentiation in mature CD4 (+) T cells and has been associated with T cell development.⁷⁹ One possible explanation of the functional impact is that processed meat could trigger a pro-tumorigenic inflammatory or immunological response,^{55,80} and loss of *GATA* genes or silencing of expression can increase CRC risk.⁸¹

Furthermore, the mechanism of the modifying effect of alcohol consumption on the association between rs9409565 at the 9q22.32/ Hippocampus Abundant Transcript-Like 1 (*HIATL1*) locus and CRC risk has not been understood. *HIATL1* is a member of the solute carrier group of membrane transport, which makes the move of substances (such as peptides, amino acids, proteins, metals, and neurotransmitters) directly into or out of cells possible.^{82,83} Gene expression analyses indicate that the variants at 9q22.32/*HIATL1* that interact with alcohol on CRC risk through genome-wide G×E interaction analyses can also impact *HIATL1* expression,⁶⁰ which suggests that alcohol may modify the effects of *HIATL1* on CRC risk through its influence on *HIATL1* expression levels.⁶⁰

Challenges for G×E interaction studies

Studies of G×E interactions require much larger sample sizes than main effect analyses due to small effect sizes, multiple testing, misclassification due to imperfect measures of environmental exposures and more model parameters.^{5,84–86} Even if the sample size is large enough to detect interactions with common exposures, it may still be insufficient when analyzing relatively rare exposures or genotypes of interest.⁵ Meta-analyses of existing G×E interactions can address this sample size limitation, though investigators should be aware of other issues, such as i) inconsistencies in the definitions of exposures and outcomes and ii) differences in study designs, tools for assessing exposures, distributions of exposures, statistical analyses, presentation of results and publication bias.⁸⁷ Additionally, the scale and distribution of environmental exposures in a population can also influence power. In dietary studies, it is suggested that the diet under investigation should be sufficiently variable in the population to allow evaluation of various intakes, and risk should be reported per similar units of exposure (e.g. per 100 g meat intake per day) to ensure that comparisons between populations is possible.²¹ In studies performed under the framework of GWAS consortia, the nature of exposure assessment may be different from other studies. Thus, harmonization of exposure may have been more difficult. Consequently, an increased sample size and decreased quality of exposure data, as well as a fully validated design may help to address the measurement error issue.⁵

Strengths and limitations of current review

The strengths of umbrella reviews have been described elsewhere.^{13,88,89} In our study, we found moderate evidence for some G×E interactions on CRC risk, though most of these interaction effects were tenuous due to the lack of main genetic effects and/or environmental effects. The proposed biological mechanisms for the G×E interactions are hypothetical and in the absence of experimental studies could not be used to prove causality. Ideally, evidence from a model system or/and organism with genetic variations in the gene/polymorphisms of interest and exposed to the physiological dose of the environmental factor (aspirin, processed meat, alcohol, and sex hormone) is required to support the epidemiological evidence described here. Thus, further replication and functional studies are required to confirm our findings and understand the biologic implications of the interactions.

Our study has limitations. First, interaction effects that have not yet been assessed through meta-analyses or systematic reviews would not have been included, since umbrella reviews do not focus on individual studies. Second, we did not use an established tool to assess the risk of bias in the included observational studies, because available tools such as Q-Genie⁹⁰ and the Newcastle-Ottawa Scale do not capture aspects relevant to the G×E assessment. Third, we excluded reviews without explicit systematic literature searches in order to avoid bias, but this could have resulted in the exclusion of syntheses of literature that have not been systematic.⁹¹ Fourth, interactions with limited evidence or limited sample size may have led to false-negative findings for some joint effects that have long been thought to exist. False-positives may also exist, although we have applied our criteria to assess the evidence to minimize biases. In our study, we used I^2 to assess the heterogeneity in the evaluation, however, it has been reported that I^2 represents what proportion of the observed variance would remain if we could eliminate the sampling error rather than how much the effect size varies.⁹² Additionally, we would miss interactions in which there were no marginal effects of genotype or exposure on CRC risk, since the Venice criteria only aim to grade the credibility of evidence of significant main effects; and 2 of the identified G×E interactions in our study could not be properly evaluated due to lack of information required to apply these criteria.

We combined systematic literature reviews, candidate and genome-wide G×E studies together. The main reason for this was to provide a comprehensive overview of the existing literature on G×E interaction in relation to CRC risk. However, an issue is that in each of these types of studies, different criteria and *p* value thresholds are typically used to evaluate the credibility of findings. Candidate gene studies with liberal *p* value thresholds might be biased toward false positive findings, while the genome-wide approach is prone to false negative observations, because true interactions may not reach a stringent genome-wide significance threshold. This, as well as other issues such as genotyping and imputation problems, could be a

reason why none of the interactions identified in the candidate gene studies were replicated in genome-wide G×E studies.

The extension of Venice criteria used here has been applied in the past to assess joint effects of environments and genes on risk of multiple cancers.^{7,10} This method does not take into account other lines of evidence, such as biological plausibility of observed associations, biological gradient of effects, coherence of the observations across multiple type of studies or support by experiments. Although other guidelines are more comprehensive in the range of evidence considered such as the newly developed “integrative research” method that combines causal criteria of Austin Bradford Hill with graphical models,⁹³ we did not use it to evaluate the evidence in our study, since some of the criteria are difficult to apply and interpret in molecular epidemiology (e.g. temporality).⁹⁴

Conclusions

Our assessment maps the status of evidence on the associations between G×E interactions and CRC risk. Despite the identified studies exploring a wide variety of G×E interactions on CRC, we conclude that we did not find highly convincing evidence for any interactions, but several associations were found to have moderate strength of evidence using our set of guidelines.

Though most of the evaluated G×E interactions in our study were with no main genetic effects, it has been suggested that such kind of risk loci that have not been readily detected in traditional GWAS may be identified by testing for interactions between SNPs and environmental risk factors, even though there is no strong evidence for a G×E interaction.^{5,6} Thus, studies incorporating accurate assessment of environmental exposures are encouraged not only to identify novel G×E interactions, but also to discover novel risk loci for CRC by characterizing any underlying G×E interactions.⁵ Moreover, there remains insufficient evidence for G×E interactions on CRC risk, and some G×E interactions without strong evidence may still be important in CRC prevention. Hence, studies with large sample sizes and further functional studies are required to identify important G×E interactions that could have public health impact, so as to shed light to CRC etiology and to allow for more specific risk assessment for early-detection or prevention strategies.

Acknowledgements

The work was supported by funding for the infrastructure and staffing of the Edinburgh CRUK Cancer Research Centre. Acknowledgments for the Scottish Colorectal Cancer Study: We thank the participants in all

of the studies that contributed to this piece of work and all the recruitment teams and collaborators who make such studies possible. We acknowledge the excellent technical support from Marion Walker. We are grateful to Ruth Wilson, Donna Markie and all those who continue to contribute to recruitment, data collection and data curation for the Study of Colorectal Cancer in Scotland studies. In addition to all consultant colorectal surgeons who provided data on their patients, we are also indebted to the chairs and offices of the managed clinical networks throughout Scotland who contributed substantially to clinicopathologic data and staging information. We acknowledge the expert support on sample preparation from the Genetics Core of the Edinburgh Wellcome Trust Clinical Research Facility.

Acknowledgements for the GECCO consortium: ASTERISK: We are very grateful to Dr. Bruno Buecher without whom this project would not have existed. We also thank all those who agreed to participate in our study, including the patients and the healthy control persons, as well as all the physicians, technicians and students. DACHS: We thank all participants and cooperating clinicians, and Ute Handte-Daub, Utz Benscheid, Muhabbet Celik and Ursula Eilber for excellent technical assistance. Harvard cohorts (HPFS, NHS, PHS): We would like to thank the participants and staff of the HPFS, NHS and PHS for their valuable contributions as well as the after state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY. The authors assume full responsibility for analyses and interpretation of these data. PLCO: The authors thank the PLCO Cancer Screening Trial screening center investigators and the staff from Information Management Services Inc and Westat Inc. Most importantly, we thank the study participants for their contributions that made our study possible. PMH: The authors would like to thank the study participants and staff of the Hormones and Colon Cancer study. WHI: The authors thank the WHI investigators and staff for their dedication, and the study participants for making the program possible. A full listing of WHI investigators can be found at: <http://www.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20Short%20List.pdf>.

Conflict of interest

None declared.

Ethical approval

Not required.

Data sharing

No additional data available.

Authors' contributions

Study design: MT and ET; Literature search: TY; Study selection: TY and MT; Data extraction: TY and ZM; Data analysis: TY and MT; Study draft and revision: TY, MT, ET, XL, ZM, JL, SMF, JPAI, MGD and HC; Article guarantor: Dr. Maria Timofeeva and Dr. Evropi Theodoratou.

References

1. Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* 2015;136:E359–86.
2. Bode AM, Dong Z, Wang H. Cancer prevention and control: alarming challenges in China. *Natl Sci Rev* 2016;3:117–27.
3. Tze CN, Fitzgerald H, Qureshi A, et al. Pioneering annual colorectal cancer screening and treatment targeting Low income communities in Malaysia (2010/2015). *Asian Pac J Cancer Prev* 2016;17:3179–83.
4. Munoz M, Pong-Wong R, Canela-Xandri O, et al. Evaluating the contribution of genetics and familial shared environment to common disease using the UKbiobank. *Nat Genet* 2016;48:980–3.
5. Hutter CM, Mechanic LE, Chatterjee N, et al. Gene-environment interactions in cancer epidemiology: a National Cancer Institute think tank report. *Genet Epidemiol* 2013;37:643–57.

6. Gauderman WJ, Zhang P, Morrison JL, et al. Finding novel genes by testing G x E interactions in a genome-wide association study. *Genet Epidemiol* 2013;37:603–13.
7. Theodoratou E, Timofeeva M, Li X, et al. Nature, nurture, and cancer risks: genetic and nutritional contributions to cancer. *Annu Rev Nutr* 2017;37: 293–320.
8. Boffetta P, Winn DM, Ioannidis JP, et al. Recommendations and proposed guidelines for assessing the cumulative evidence on joint effects of genes and environments on cancer occurrence in humans. *Int J Epidemiol* 2012;41: 686–704.
9. Ioannidis JP, Boffetta P, Little J, et al. Assessment of cumulative evidence on genetic associations: interim guidelines. *Int J Epidemiol* 2008;37:120–32.
10. World Cancer Research Fund/American Institute for Cancer Research. Diet, Nutrition, Physical Activity and Cancer: a Global Perspective 2018. <https://www.wcrf.org/dietandcancer>.
11. World Cancer Research Fund/American Institute for Cancer Research. Continuous Update Project Report: Diet, Nutrition, Physical Activity and Colorectal Cancer, 2017. <https://www.wcrf.org/sites/default/files/Colorectal-Cancer-2017-Report.pdf>.
12. World Cancer Research Fund. International Systematic Literature Review. The Associations between Food, Nutrition and Physical Activity and the Risk of Colorectal Cancer, 2016. https://www.wcrf.org/sites/default/files/CUP_colorectal_cancer_SLR_2016lo.pdf.
13. Bellou V, Belbasis L, Tzoulaki I, et al. Environmental risk factors and Parkinson's disease: an umbrella review of meta-analyses. *Parkinsonism Relat Disord* 2016;23:1–9.
14. Welter D, MacArthur J, Morales J, et al. The NHGRI GWAS catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res* 2014;42: D1001–6.
15. Beck T, Hastings RK, Gollapudi S, et al. GWAS central: a comprehensive resource for the comparison and interrogation of genome-wide association studies. *Eur J Hum Genet* 2014;22:949–52.
16. Khoury MJ, Bertram L, Boffetta P, et al. Genome-wide association studies, field synopses, and the development of the knowledge base on genetic variation and human diseases. *Am J Epidemiol* 2009;170:269–79.
17. Vineis P, Manuguerra M, Kavvoura FK, et al. A field synopsis on low-penetrance variants in DNA repair genes and cancer susceptibility. *J Natl Cancer Inst* 2009;101:24–36.
18. Dimitrakopoulou VI, Travis RC, Shui IM, et al. Interactions between genome-wide significant genetic variants and circulating concentrations of 25-Hydroxyvitamin D in relation to prostate cancer risk in the National Cancer Institute BPC3. *Am J Epidemiol* 2017;185:452–64.
19. Theodoratou E, Farrington SM, Tenesa A, et al. Dietary vitamin B6 intake and the risk of colorectal cancer. *Cancer Epidemiol Biomarkers Prev* 2008;17:171–82.
20. Peters U, Jiao S, Schumacher FR, et al. Identification of genetic susceptibility loci for colorectal tumors in a genome-wide meta-analysis. *Gastroenterology* 2013;144:799–807.
21. Andersen V, Holst R, Vogel U. Systematic review: diet-gene interactions and the risk of colorectal cancer. *Aliment Pharmacol Ther* 2013;37:383–91.
22. Andersen V, Vogel U. Systematic review: interactions between aspirin, and other nonsteroidal anti-inflammatory drugs, and polymorphisms in relation to colorectal cancer. *Aliment Pharmacol Ther* 2014;40:147–59.
23. Andersen V, Vogel U. Interactions between meat intake and genetic variation in relation to colorectal cancer. *Genes Nutr* 2015;10:448.
24. Corella D, Ordovas JM. Interactions between dietary n-3 fatty acids and genetic variants and risk of disease. *Br J Nutr* 2012;107:S271–S83.
25. Cornelis MC. Gene-coffee interactions and health. *Curr Nutr Rep* 2014;3:178–95.
26. Cross JT, Poole EM, Ulrich CM. A review of gene-drug interactions for nonsteroidal anti-inflammatory drug use in preventing colorectal neoplasia. *Pharmacogenomics J* 2008;8: 237–47.
27. Eichholzer M, Luthy J, Moser U, et al. Folate and the risk of colorectal, breast and cervix cancer: the epidemiological evidence. *Swiss Med Wkly* 2001; 131:539–49.
28. Houlston RS, Tomlinson IP. Polymorphisms and colorectal tumor risk. *Gastroenterology* 2001;121: 282–301.
29. Karahalil B, Bohr VA, Wilson DM 3rd. Impact of DNA polymorphisms in key DNA base excision repair proteins on cancer risk. *Hum Exp Toxicol* 2012;31:981–1005.
30. Klarich DS, Brasser SM, Hong MY. Moderate alcohol consumption and colorectal cancer risk. *Alcohol Clin Exp Res* 2015;39:1280–91.
31. Kostner K, Denzer N, Muller CS, et al. The relevance of vitamin D receptor (VDR) gene polymorphisms for cancer: a review of the literature. *Anticancer Res* 2009;29:3511–36.
32. Lenihan-Geels G, Bishop KS, Ferguson LR. Cancer risk and eicosanoid production: interaction between the protective effect of long chain Omega-3 polyunsaturated fatty acid intake and genotype. *J Clin Med* 2016;5(2):pii:E25.
33. Sharp L, Little J. Polymorphisms in genes involved in folate metabolism and colorectal neoplasia: a HuGE review. *Am J Epidemiol* 2004;159:423–43.
34. Shin A, Kim J. Effect modification of meat intake by genetic polymorphisms on colorectal neoplasia susceptibility. *Asian Pac J Cancer Prev* 2010;11: 281–7.
35. Simonds NI, Ghazarian AA, Pimentel CB, et al. Review of the gene-environment interaction literature in cancer: what do we know? *Genet Epidemiol* 2016;40:356–65.
36. Ananthakrishnan AN, Du M, Berndt SI, et al. Red meat intake, NAT2, and risk of colorectal cancer: a pooled analysis of 11 studies. *Cancer Epidemiol Biomarkers Prev* 2015;24:198–205.
37. Ding W, Zhou DL, Jiang X, et al. Methionine synthase A2756G polymorphism and risk of colorectal adenoma and cancer: evidence based on 27 studies. *PLoS One* 2013;8:e60508.
38. Hiraki LT, Joshi AD, Ng K, et al. Joint effects of colorectal cancer susceptibility loci, circulating 25-hydroxyvitamin D and risk of colorectal cancer. *PLoS One* 2014;9:e92212.
39. Hutter CM, Chang-Claude J, Slattery ML, et al. Characterization of gene-environment interactions for colorectal cancer susceptibility loci. *Cancer Res* 2012;72:2036–44.
40. Hutter CM, Slattery ML, Duggan DJ, et al. Characterization of the association between 8q24 and colon cancer: gene-environment exploration and meta-analysis. *BMC Cancer* 2010; 10:670.
41. Jiao S, Hsu L, Bezieau S, et al. SBERIA: set-based gene-environment interaction test for rare and common variants in complex diseases. *Genet Epidemiol* 2013;37:452–64.
42. Kantor ED, Hutter CM, Minnier J, et al. Gene-environment interaction involving recently identified colorectal cancer susceptibility loci. *Cancer Epidemiol Biomarkers Prev* 2014;23:1824–33.
43. Liu Y, Qin H, Zhang Y, et al. P53 codon 72 polymorphism and colorectal cancer: a meta-analysis of epidemiological studies. *Hepatogastroenterology* 2011;58:1926–9.
44. Nan H, Morikawa T, Suuriniemi M, et al. Aspirin use, 8q24 single nucleotide polymorphism rs6983267, and colorectal cancer according to CTNNB1 alterations. *J Natl Cancer Inst* 2013;105: 1852–61.
45. Pabalan N, Bapat B, Sung L, et al. Cyclin D1 Pro241Pro (CCND1-G870A) polymorphism is associated with increased cancer risk in human populations: a meta-analysis. *Cancer Epidemiol Biomarkers Prev* 2008;17:2773–81.
46. Pabalan N, Jarjanazi H, Ozcelik H. A meta-analysis of the C1420T polymorphism in cytosolic serine hydroxymethyltransferase (SHMT1) among Caucasian colorectal cancer populations. *Int J Colorectal Dis* 2013;28:925–32.
47. Raimondi S, Botteri E, Iodice S, et al. Gene-smoking interaction on colorectal adenoma and cancer risk: review and meta-analysis. *Mutat Res* 2009;670:6–14.
48. Smits KM, Gaspari L, Weijenberg MP, et al. Interaction between smoking, GSTM1 deletion and colorectal cancer: results from the GSEC study. *Biomarkers* 2003;8:299–310.
49. Tse G, Eslick GD. Cruciferous vegetables and risk of colorectal neoplasms: a systematic review and meta-analysis. *Nutr Cancer* 2014;66:128–39.
50. Wan H, Zhou Y, Yang P, et al. Genetic polymorphism of glutathione S-transferase T1 and the risk of colorectal cancer: a meta-analysis. *Cancer Epidemiol* 2010;34:66–72.
51. Wang H, Iwasaki M, Haiman CA, et al. Interaction between red meat intake and NAT2 genotype in increasing the risk of colorectal cancer in Japanese and African Americans. *PLoS One* 2015;10:e0144955.
52. Wang X, Xing GH, Fan CC. Association between the FAS rs2234767G/a polymorphism and cancer risk: a systematic review and meta-analysis. *DNA Cell Biol* 2014;33:320–7.
53. Lq Z, Jn Z, Wang J, et al. Absence of association between N-Acetyltransferase 2 acetylator status and colorectal cancer susceptibility: based on evidence from 40 studies. *PLoS One* 2012;7(3): e32425.
54. Danfeng S, Xia W, Jingyuan F. Meta-analysis of the relationship between methylenetetrahydrofolate reductase gene polymorphism and colon cancer. *Gastroenterology* 2006;11:516–21.
55. Figueiredo JC, Hsu L, Hutter CM, et al. Genome-wide diet-gene interaction analyses for risk of colorectal cancer. *PLoS Genet* 2014;10:e1004228.
56. Figueiredo JC, Lewinger JP, Song C, et al. Genotype-environment interactions in microsatellite stable/microsatellite instability-low colorectal cancer: results from a genome-wide association study. *Cancer Epidemiol Biomarkers Prev* 2011;20:758–66.

57. Nan H, Hutter CM, Lin Y, et al. Association of aspirin and NSAID use with risk of colorectal cancer according to genetic variants. *JAMA* 2015; 313:1133–42.
58. Du M, Zhang X, Hoffmeister M, et al. No evidence of gene-calcium interactions from genome-wide analysis of colorectal cancer risk. *Cancer Epidemiol Biomarkers Prev* 2014;23: 2971–6.
59. Garcia-Albeniz X, Rudolph A, Hutter C, et al. CYP24A1 variant modifies the association between use of oestrogen plus progestogen therapy and colorectal cancer risk. *Br J Cancer* 2016; 114:221–9.
60. Gong J, Hutter CM, Newcomb PA, et al. Genome-wide interaction analyses between genetic variants and alcohol consumption and smoking for risk of colorectal cancer. *PLoS Genet* 2016;12:e1006296.
61. Jiao S, Peters U, Berndt S, et al. Powerful set-based gene-environment interaction testing framework for complex diseases. *Genet Epidemiol* 2015;39:609–18.
62. Siegert S, Hampe J, Schafmayer C, et al. Genome-wide investigation of gene-environment interactions in colorectal cancer. *Hum Genet* 2013;132: 219–31.
63. Qiao Y, Yang T, Gan Y, et al. Associations between aspirin use and the risk of cancers: a meta-analysis of observational studies. *BMC Cancer* 2018;18:288.
64. Tanskanen T, van den Berg L, Valimaki N, et al. Genome-wide association study and meta-analysis in northern European populations replicate multiple colorectal cancer risk loci. *Int J Cancer* 2017;142(3):540–546.
65. Orlando G, Law PJ, Palin K, et al. Variation at 2q35 (PNKD and TMBIM1) influences colorectal cancer risk and identifies a pleiotropic effect with inflammatory bowel disease. *Hum Mol Genet* 2016;25:2349–59.
66. Lin KJ, Cheung WY, Lai JY, et al. The effect of estrogen vs combined estrogen-progestogen therapy on the risk of colorectal cancer. *Int J Cancer* 2012;130:419–30.
67. Schumacher FR, Schmit SL, Jiao S, et al. Genome-wide association study of colorectal cancer identifies six new susceptibility loci. *Nat Commun* 2015;6:7138.
68. Pomerantz MM, Ahmadiyeh N, Jia L, et al. The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer. *Nat Genet* 2009;41:882–4.
69. Tuupanen S, Turunen M, Lehtonen R, et al. The common colorectal cancer predisposition SNP rs6983267 at chromosome 8q24 confers potential to enhanced Wnt signaling. *Nat Genet* 2009;41: 885–90.
70. Sur IK, Hallikas O, Vaharautio A, et al. Mice lacking a Myc enhancer that includes human SNP rs6983267 are resistant to intestinal tumors. *Science (New York, NY)* 2012;338:1360–3.
71. Bos CL, Kodach LL, van den Brink GR, et al. Effect of aspirin on the Wnt/beta-catenin pathway is mediated via protein phosphatase 2A. *Oncogene* 2006;25:6447–56.
72. Fresno Vara JA, Casado E, de Castro J, et al. PI3K/Akt signalling pathway and cancer. *Cancer Treat Rev* 2004;30:193–204.
73. Kaur J, Sanyal SN. PI3-kinase/Wnt association mediates COX-2/PGE(2) pathway to inhibit apoptosis in early stages of colon carcinogenesis: chemoprevention by diclofenac. *Tumour Biol* 2010; 31:623–31.
74. Grivnenkov SI, Karin M. Inflammatory cytokines in cancer: tumour necrosis factor and interleukin 6 take the stage. *Ann Rheum Dis* 2011;70(Suppl 1):i104–8.
75. Klampfer L. Cytokines, inflammation and colon cancer. *Curr Cancer Drug Targets* 2011;11:451–64.
76. Symmons O, Uslu VV, Tsujimura T, et al. Functional and topological characteristics of mammalian regulatory domains. *Genome Res* 2014;24:390–400.
77. Bareis P, Bises G, Bischof MG, et al. 25-hydroxy-vitamin d metabolism in human colon cancer cells during tumor progression. *Biochem Biophys Res Commun* 2001;285:1012–7.
78. Dong LM, Ulrich CM, Hsu L, et al. Vitamin D related genes, CYP24A1 and CYP27B1, and colon cancer risk. *Cancer Epidemiol Biomarkers Prev* 2009;18:2540–8.
79. Hosoya T, Maillard I, Engel JD. From the cradle to the grave: activities of GATA-3 throughout T-cell development and differentiation. *Immunol Rev* 2010;238:110–25.
80. Hedlund M, Padler-Karavani V, Varki NM, et al. Evidence for a human-specific mechanism for diet and antibody-mediated inflammation in carcinoma progression. *Proc Natl Acad Sci U S A* 2008;105:18936–41.
81. Zheng R, Blobel GA. GATA transcription factors and cancer. *Genes Cancer* 2010;1:1178–88.
82. Schlessinger A, Matsson P, Shima JE, et al. Comparison of human solute carriers. *Protein Sci* 2010;19:412–28.
83. Hoglund PJ, Nordstrom KJ, Schioth HB, et al. The solute carrier families have a remarkably long evolutionary history with the majority of the human families present before divergence of Bilaterian species. *Mol Biol Evol* 2011;28:1531–41.
84. Smith PG, Day NE. The design of case-control studies: the influence of confounding and interaction effects. *Int J Epidemiol* 1984;13:356–65.
85. Thomas D. Gene-environment-wide association studies: emerging approaches. *Nat Rev Genet* 2010;11:259–72.
86. Burton PR, Hansell AL, Fortier I, et al. Size matters: just how BIG is BIG?: quantifying realistic sample size requirements for human genome epidemiology. *Int J Epidemiol* 2009;38: 263–73.
87. Cornelis MC, Hu FB. Gene-environment interactions in the development of type 2 diabetes: recent progress and continuing challenges. *Annu Rev Nutr* 2012;32:245–59.
88. Theodoratou E, Tzoulaki I, Zgaga L, et al. Vitamin D and multiple health outcomes: umbrella review of systematic reviews and meta-analyses of observational studies and randomised trials. *BMJ (Clin Res Ed)* 2014;348:g2035.
89. Belbasis L, Savvidou MD, Kanu C, et al. Birth weight in relation to health and disease in later life: an umbrella review of systematic reviews and meta-analyses. *BMC Med* 2016;14:147.
90. Sohani ZN, Meyre D, de Souza RJ, et al. Assessing the quality of published genetic association studies in meta-analyses: the quality of genetic studies (Q-genie) tool. *BMC Genet* 2015;16:50.
91. Li X, Meng X, Timofeeva M, et al. Serum uric acid levels and multiple health outcomes: umbrella review of evidence from observational studies, randomised controlled trials, and Mendelian randomisation studies. *BMJ (Clin Res Ed)* 2017;357: j2376.
92. Borenstein M, Higgins JP, Hedges LV, et al. Basics of meta-analysis: I(2) is not an absolute measure of heterogeneity. *Res Synth Methods* 2017;8:5–18.
93. Geneletti S, Gallo V, Porta M, et al. Assessing causal relationships in genomics: from Bradford-Hill criteria to complex gene-environment interactions and directed acyclic graphs. *Emerg Themes Epidemiol* 2011;8:5.
94. Fedak KM, Bernal A, Capshaw ZA, et al. Applying the Bradford Hill criteria in the 21st century: how data integration has changed causal inference in molecular epidemiology. *Emerg Themes Epidemiol* 2015;12:14.